Environmental Research 116 (2012) 1–10

Contents lists available at SciVerse ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation

Mercedes A. Bravo [a,*], Montserrat Fuentes [b,1], Yang Zhang [c,2], Michael J. Burr [c], Michelle L. Bell [a,3]

[a] School of Forestry and Environmental Studies, Yale University, 195 Prospect St., New Haven, CT 06511, USA
[b] Statistics Department, North Carolina State University, 2311 Stinson Drive, Campus Box 8203, Raleigh, NC 27695-8203, USA
[c] Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, 2800 Faucette Drive, 1125 Jordan Hall, Raleigh, NC 27695-8208, USA

ARTICLE INFO

Article history:
Received 31 August 2011
Received in revised form
10 April 2012
Accepted 18 April 2012
Available online 10 May 2012

Keywords:
$PM_{2.5}$
Ozone
Community multiscale air quality model
CMAQ
Exposure assessment
Health assessment

ABSTRACT

Air quality modeling could potentially improve exposure estimates for use in epidemiological studies. We investigated this application of air quality modeling by estimating location-specific (point) and spatially-aggregated (county level) exposure concentrations of particulate matter with an aerodynamic diameter less than or equal to 2.5 μm ($PM_{2.5}$) and ozone (O_3) for the eastern U.S. in 2002 using the Community Multi-scale Air Quality (CMAQ) modeling system and a traditional approach using ambient monitors. The monitoring approach produced estimates for 370 and 454 counties for $PM_{2.5}$ and O_3, respectively. Modeled estimates included 1861 counties, covering 50% more population. The population uncovered by monitors differed from those near monitors (e.g., urbanicity, race, education, age, unemployment, income, modeled pollutant levels). CMAQ overestimated O_3 (annual normalized mean bias=4.30%), while modeled $PM_{2.5}$ had an annual normalized mean bias of −2.09%, although bias varied seasonally, from 32% in November to −27% in July. Epidemiology may benefit from air quality modeling, with improved spatial and temporal resolution and the ability to study populations far from monitors that may differ from those near monitors. However, model performance varied by measure of performance, season, and location. Thus, the appropriateness of using such modeled exposures in health studies depends on the pollutant and metric of concern, acceptable level of uncertainty, population of interest, study design, and other factors.

1. Introduction

Exposure estimates in air pollution and health studies are commonly assessed using data from ambient air quality monitors. Many places, particularly urban areas, have established monitoring networks with historical, publicly available data. Several methods for estimating exposures to air pollutants exist, including monitor-based approaches such as proximity-based assessments and statistical interpolation, as well as land-use regression and air quality modeling (Jerrett et al., 2005a). Utilizing data from existing monitoring networks remains popular, due to cost considerations, data availability, and population coverage.

Most epidemiological studies of air pollution are based in urban areas, and most monitors for criteria pollutants, such as particulate matter with an aerodynamic diameter less than or equal to 2.5 μm ($PM_{2.5}$) and ozone (O_3), are located in areas with a high percentage of the population living in urban and suburban environments (Bell, 2006). Monitoring data may be best at estimating exposure for populations close to the monitor's location (Sarnat et al., 2006), who are disproportionately in urban environments, with less spatial coverage for exposure estimation in rural environments. Characteristics of air pollution (e.g., chemical components, particle properties) vary spatially (Bell et al., 2007a) and may differ between areas near and far from monitors.

Ambient monitors offer limited temporal resolution and coverage; many do not operate continuously throughout the year. In the U.S., most $PM_{2.5}$ monitors record a 24-hour measurement every three days, with some monitors sampling daily or every six days. Ozone is usually measured hourly, but only for a portion of the year (e.g., April–September). Limited spatial and temporal resolution hinders statistical power and determines the types of scientific questions that can be investigated, leaving questions about health effects of cumulative exposures and in rural environments.

One approach to address these limitations is application of three-dimensional (3-D) air quality models such as the Community

* Corresponding author. Fax: +203 436 9135.
E-mail addresses: mercedes.bravo@yale.edu (M.A. Bravo), fuentes@stat.ncsu.edu (M. Fuentes), yang_zhang@ncsu.edu (Y. Zhang), mikeburr3@gmail.com (M.J. Burr), michelle.bell@yale.edu (M.L. Bell).
[1] Fax: +919 515 7591.
[2] Fax: +925 515 7802.
[3] Fax: +203 436 9135.

0013-9351/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.envres.2012.04.008

Multi-Scale Air Quality (CMAQ) modeling system. CMAQ is a sophisticated, state-of-the-art, regional air quality model capable of estimating concentrations of multiple pollutants at local, regional, or continental scales (Byun and Schere, 2006). CMAQ combines input from a meteorological model and an emissions model with simulation of chemical and physical processes to describe pollutant transformation, transport, and fate (Fig. S1, Supplementary Material). Output includes gridded estimates of pollutant concentrations and deposition fluxes. Compared to approaches relying exclusively on monitor data, the use of CMAQ results to estimate exposures offers improved spatial coverage and greater spatial and temporal resolution.

This study evaluates use of regional air quality modeling results, using CMAQ as an example, for generating estimates of exposure to air pollutants, as an alternative or supplement to monitoring data. The primary objectives of this analysis are to evaluate limitations and advantages of using CMAQ to estimate exposure levels. To achieve this we: (1) conducted an evaluation of CMAQ performance, emphasizing the use of model results for exposure estimates; (2) compared characteristics of populations covered and not covered by the monitoring network; and (3) generated and compared location-specific and spatially-aggregated exposure estimates using monitoring data and modeling results.

2. Data and methods

2.1. Air pollution data

We used results from a simulation of CMAQ version 4.5.1 covering much of the eastern U.S. at a 12 km horizontal grid resolution. Initial and boundary conditions for meteorology and chemistry were extracted from a 36 km simulation conducted by the Visibility State and Tribal Association of the Southeast. Meteorology was simulated using the Pennsylvania State University/National Center for Atmospheric Research 5th generation mesocscale model version 3.7. The emissions inventory was based on the 1999 National Emissions Inventory version 2, and was processed using the Sparse Matrix Operator Kernel Emissions version 2.1. Ground-level $PM_{2.5}$ (24-hour average) and O_3 (8-hour maximum) concentrations were simulated for grid cells (12×12 km) for each day in 2002. Each estimate represents a volume-averaged concentration over the grid cell. Information on the CMAQ system (Byun and Schere, 2006; Zhang et al., 2006b, 2006c) and additional details on this specific simulation, including meteorological and emissions data, are provided elsewhere (Burr and Zhang, 2011; Morris et al., 2009; Olerud and Sims, 2004; Queen and Zhang, 2008).

Methods for estimating exposure from monitoring data were designed to emulate those in the epidemiologic literature (Miller et al., 2007; Peng et al., 2008; Pope et al., 2002; Sarnat et al., 2009). Monitoring data were obtained from the Air Quality System, which contains data collected by the U.S. Environmental Protection Agency (EPA), state, local, and tribal air pollution control agencies. Monitoring data were daily 24-hour average $PM_{2.5}$ and daily maximum 8-hour O_3 levels for 2002, metrics by which $PM_{2.5}$ and O_3 are regulated. Most $PM_{2.5}$ monitors provide data every three days, with some sampling every day or every six days. For most monitors, O_3 was measured every day during "O_3 season" (typically April–September). Only Federal Reference Method-compliant $PM_{2.5}$ monitors and Federal Equivalence Method-compliant O_3 monitors were considered.

We omitted observations the U.S.EPA coded as problematic (e.g., "lab issues"). Of 738 U.S.EPA $PM_{2.5}$ monitoring sites in the study area (defined as a monitor within the CMAQ domain or outside the domain but within 100 km of its border), 117 (16%) had multiple monitors. Most sites with co-located monitors had only two monitors, for a total of 857 monitors in the study area. Of 752 U.S.EPA O_3 monitoring sites in the study area, only one had multiple monitors, for a total of 753 O_3 monitors. Co-located monitors were treated as repeated measurements, and averaged for each day.

In the analysis of demographic characteristics of populations in counties with and without monitors, no monitors were omitted due to data availability. For exposure analysis, monitors with insufficient data to meet inclusion criteria were excluded to reflect exposure estimation methods typical of health effects studies. Inclusion criteria were selected with the purpose of being sufficiently stringent to avoid over-representation of particular seasons. For $PM_{2.5}$, inclusion criteria were developed based on a one-in-three-day sampling frequency. Thus, complete data for a monitor in 2002 would include 121 days. Monitors with less than 76% complete data (i.e., fewer than 91 observations) were excluded. To ensure seasonal representation, the year was divided into 13 periods of 28 days. Monitors were excluded if they had fewer than 11 (of 13) 28-day periods with at least one observation per week for three or more weeks in the period. Within the study area, 218 (25%) of 857 $PM_{2.5}$ monitors did not meet inclusion criteria.

Inclusion criteria for O_3 were based on a daily measurement frequency during April–September. We included only monitors with a daily 8-hour maximum reported for a minimum of 75% of days in April through September. Monitors were also required to have data for at least 50% of days in each month for 5 or more months of the 6-month O_3 season. Within the study area, 729 (97%) of 753 O_3 monitors met inclusion criteria for exposure analysis. Fig. S2, Supplementary Material provides the CMAQ domain and monitor locations.

2.2. Model evaluation: model results vs. monitor data

The CMAQ model has been extensively assessed and updated based on evaluation results, review panels, and improvements in understanding of modeled processes. Evaluations indicate CMAQ generally provides reasonable pollution estimates and also identify specific conditions, locations or processes in which performance could be improved (Appel et al., 2007; Bailey et al., 2007; Baker and Scheff, 2007; Boylan and Russell, 2006; Eder et al., 2006; Eder and Yu, 2006; Mueller, 2009; Phillips and Finkelstein, 2006b; Swall and Davis, 2006; Tesche et al., 2006; Zhang et al., 2006b, 2006c). CMAQ was originally designed for purposes of policy evaluation and assessing attainment of air quality standards. Thus, previous evaluations were conducted to assess whether the model adequately performs those functions. Our evaluation was conducted to identify systematic biases for this specific simulation that could impact exposure estimates in health studies. Model results in the form of grid cell concentrations were compared with observations at monitors within the grid cell's boundaries. Monitoring data used in the model evaluation were equivalent to the monitoring data used to derive exposure estimates (i.e., subject to the same inclusion/completeness criteria).

Previous studies compared monitored levels of $PM_{2.5}$ and O_3 to CMAQ estimates for the eastern, central, and contiguous U.S. (Baker and Scheff, 2007; Boylan and Russell, 2006; Eder and Yu, 2006; Zhang et al., 2006b, 2006c) using a number of metrics. The metrics used in this study, such as normalized mean bias and error and mean fractional bias and error, are frequently used in the model evaluation literature (Boylan and Russell, 2006; Eder and Yu, 2006; Zhang et al., 2006a). In addition, we considered correlation, mean bias, and root mean-square error. For all metrics other than correlation, superior model performance is indicated by values approaching zero. Formulas and description of metrics are provided in Table S1, Supplementary Material.

2.3. Population characteristics in relation to monitor locations

We investigated whether demographics of populations in locations with $PM_{2.5}$ and O_3 monitors differ from populations in areas without monitors using a suite of variables used previously as indicators of socio-economic status, racial composition, urbanicity, and other factors (Bell and Dominci, 2008; O'Neill et al., 2003b). The following variables, reported by county and obtained from the 2000 Census for all counties in the model domain, were utilized: population self-identified as African-American (Census 2000 Summary File 1, Table P3 [SF1.P3]), population living in urban settings (Census 2000 Summary File 3, Table P5 [SF3.P5]), population age 65 years and older (SF3.P8), population age 5 years and younger (SF3.P8), population using public transport (SF3.P30), population age 25 years and older with bachelor's degree (SF3.P37), population age 25 years and older with high school diploma (SF3.P37), population age 16 years and older that is unemployed (SF3.P43), median household income in 1999 (SF3.P53), and population in poverty (SF3.P87). Counties were grouped based on whether the county contained a $PM_{2.5}$ (or O_3) monitor. Inclusion criteria relating to completeness of monitoring data were not applied.

2.4. Exposure estimates

In health studies, air pollution exposure estimates can be based on a geographic area, (e.g., county, zip code), or on an individual location (e.g., study subject's residence). We generated spatially-aggregated and location-specific exposure estimates using both monitor data and CMAQ simulation results. Monitoring data represent a specific point, while simulation results are an average concentration over the grid cell volume. Thus, deriving exposure estimates from monitoring data and model results required different methods.

2.4.1. Spatially aggregated exposure estimates

We generated exposure estimates at the county level, a spatial unit commonly used in epidemiological studies (Bell et al., 2004a, 2007b; Dominici et al., 2006; Holloman et al., 2004; Janes et al., 2007; Pope et al., 2009). County level $PM_{2.5}$ and O_3 concentrations were estimated using two methods: (1) concentrations from a monitor or average of monitors located within a county; and (2) an area-weighted average of 12×12 km gridded CMAQ model results. For monitoring data, spatially-aggregated exposure estimates were generated only for counties with monitors and days with observations. Multiple monitor measurements for the same day and county were averaged. County level exposure estimates for $PM_{2.5}$ and O_3 derived from monitor data were possible only for some days in 2002 and for a subset of counties (~25%) within the model domain. County-level averages based on monitoring data incorporated estimates from all monitors in a given county, including monitors within 100 km of the study domain in order to account for counties were partially in and partially out of the study domain. Exposures based on monitoring data were not estimated for counties without monitors.

Exposure estimates from model outputs were generated for all counties with more than 98% of county area within the CMAQ domain. County level exposure estimates were calculated from an area weighted average of CMAQ grid cell(s) containing any portion of the county. County level exposure estimates derived from model results are available for all days in 2002 and all 1861 counties within the model domain.

2.4.2. *Location-specific exposure estimates*

We also generated location-specific exposure estimates reflecting pollutant concentrations at a particular point independent of political boundaries using modeling results and monitoring data. This is intended to mirror exposure assessment methods commonly used in epidemiological studies with individual level location information, such as a cohort study (Brauer et al., 2008; Jerrett et al., 2005b; Ritz et al., 2002). Many methods to generate location-specific exposures exist (e.g., inverse distance weighting, kriging). Location-specific exposure fields for $PM_{2.5}$ and O_3 were estimated using two methods: (1) using monitor data, all locations within the study area were assigned the concentration level recorded at the nearest monitor location within 50 km and were not assigned exposures if the nearest monitor was greater than 50 km away; and (2) using CMAQ results, concentration estimates in the grid cells were designated as the exposure fields. A distance of 50 km was chosen because it represented a reasonable distance for extrapolation of observed air pollutant concentrations and has been used previously in epidemiological settings (Hanigan et al., 2006; Lipsett et al., 2011; O'Donnell et al., 2011; Spencer-Hwang et al., 2011), but other distances could have been selected with similar justification.

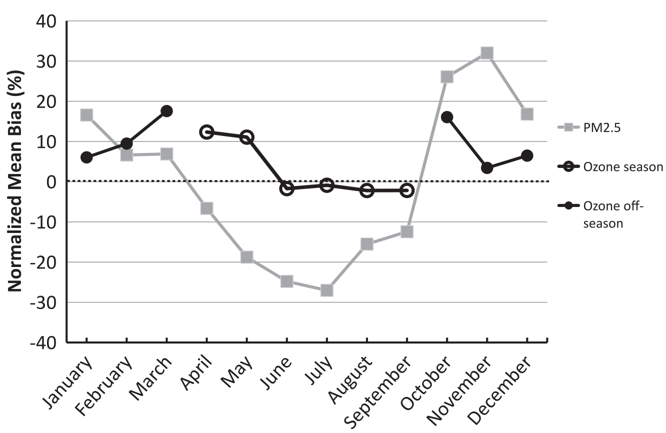

Fig. 1. Monthly normalized mean bias in simulated concentrations of $PM_{2.5}$ and O_3

3. Results

3.1. *Model evaluation*

Overall $PM_{2.5}$ concentrations, averaged across the study period and spatial domain for locations and days with both monitoring and modeled estimates were similar: 13.1 $\mu g/m^3$ for modeled concentrations and 13.4 $\mu g/m^3$ for monitor values. Maximum 8-h ozone levels were slightly higher for modeled estimates (47.5 parts per billion (ppb)]) relative to measured values (45.0 ppb). Overall, positive values for annual average bias metrics (normalized mean bias, mean bias) indicate the model overestimates O_3 levels (normalized mean bias = 4.30%, mean bias = 2.41 ppb), while negative values for $PM_{2.5}$ suggests the model tends to underestimate observed $PM_{2.5}$ (normalized mean bias = −2.09%, mean bias = −0.280 $\mu g/m^3$) (Table S2, Supplementary Material).

To identify seasonal and temporal trends in model performance, monthly normalized mean bias values were calculated (Fig. 1). Monthly normalized mean bias for O_3 ranged from −2 to 12% during the O_3 season; monthly normalized mean bias is positive for colder months when many O_3 monitors are not operated. Average monthly normalized mean bias for $PM_{2.5}$ (range: ±30%) demonstrates a distinct seasonal trend: colder months have a positive bias, while warmer months show a negative bias. Annual measures of bias may be low for $PM_{2.5}$ because seasonal trends in bias "cancel out." The mean annual correlation coefficient between simulated and observed values was 0.640 for $PM_{2.5}$ and 0.801 O_3 (correlation coefficient during O_3 season was 0.755) (Table S3 and Fig. S3 (a) and (b), Supplementary Material).

We also considered spatial trends in model performance. Annual average normalized mean bias and correlation were plotted by monitor location to evaluate whether bias and correlation differed across the study area (Figs. 2 and 3). Generally, annual average normalized mean bias for $PM_{2.5}$ were lowest (e.g., less than ±10%) in the Midwest, western Gulf coast, and northeast. Larger positive biases

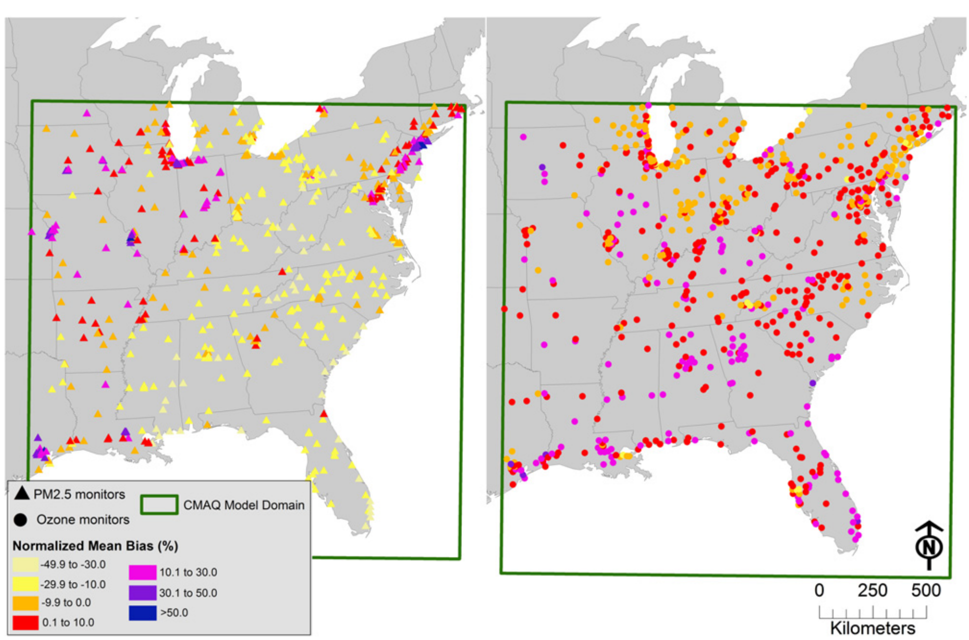

Fig. 2. Annual average normalized mean bias (by monitor location)

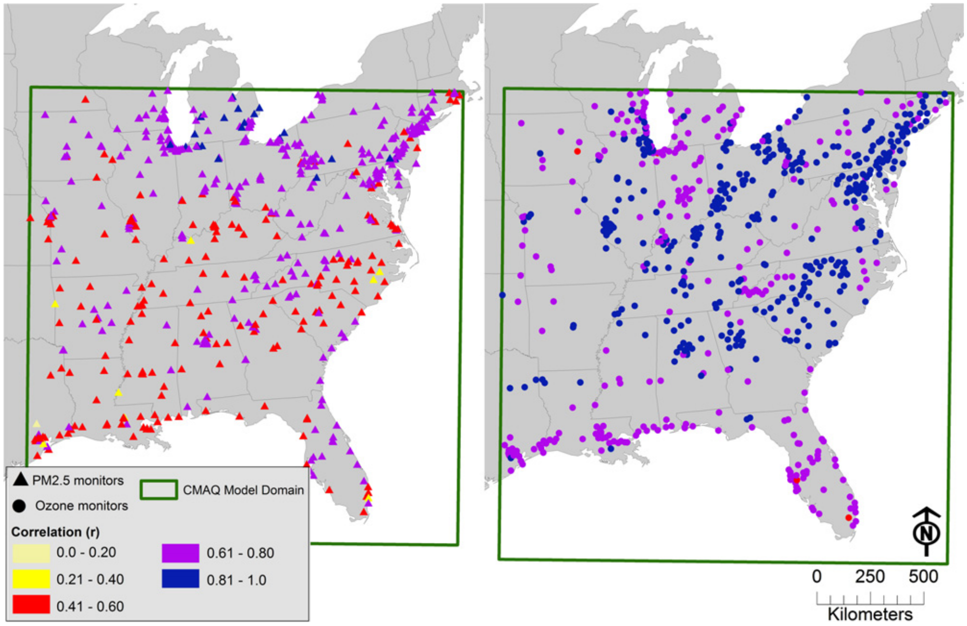

Fig. 3. Annual average correlation between observed and simulated concentrations (by monitor location)

(+10 to +30%) were concentrated in the Northeast, western and northern Midwestern states, and the Texas coast; larger negative biases (−10 to −30%) were primarily found in the eastern half of the study area. For O_3, higher annual average normalized mean bias (+10 to +30%) values were most prevalent in the Southeast, particularly in coastal areas. Generally, correlations between monitored and modeled concentrations were higher for O_3 compared to $PM_{2.5}$. Correlations for $PM_{2.5}$ tended to be highest (e.g., 0.61–0.80) in the Northeast and northern Midwest, and lower (e.g., 0.41–0.60) in parts of the Southeast and Gulf Coast. For O_3, correlations were highest (greater than 0.80) in the upper Southeast, Northeast, and Ohio River valley; correlations were consistently lower (e.g., 0.61–0.80) in Florida, the Gulf Coast, and around the Great Lakes.

3.2. Population characteristics for areas with and without monitors

Populations near monitors differed from populations farther from monitors: counties with monitors tended to have a higher percentage of individuals living in urban areas than counties without monitors, at 71.2 versus 32.9% urbanicity for $PM_{2.5}$, and 65.2 compared to 33.6% urbanicity for O_3 (Table 1). A larger proportion of individuals use public transport in counties with monitors. Counties with monitors had higher indicators of socio-economic conditions with a higher percentage of college graduates, higher median income, and a lower percentage of residents in poverty than counties without monitors. However, counties with $PM_{2.5}$ or O_3 monitors also had a lower percentage of residents with high school education than counties without monitors. Finally, counties with monitors exhibited significantly higher modeled levels of $PM_{2.5}$ and O_3 than counties without monitors (p-value less than 0.05), although actual differences between modeled annual average concentrations in counties with and without monitors was $\sim$1.5 $\mu g/m^3$ for $PM_{2.5}$ and less than 1 ppb for O_3.

3.3. Exposure estimates

3.3.1. Spatially aggregated estimates

We generated daily and annual exposure estimates aggregated at the county level for $PM_{2.5}$ and O_3 using modeling output and ambient monitors. There are 1861 counties for which more than 98% of the county area falls within the CMAQ model domain. Of these, 370 and 454 counties contained at least one monitor meeting inclusion criteria for $PM_{2.5}$ or O_3, respectively.

Table 2 describes county level exposure estimates derived from monitoring data and simulation results for $PM_{2.5}$ and O_3. The table provides information on the population covered by the exposures, the number of observations or simulation results available, and the land area covered. Exposures are provided as summary statistics (mean, standard deviation, minimum, and maximum) of annual average and daily average concentrations. This table makes comparisons between county level exposure estimates from monitor data and model results in three ways: (1) all estimates based on monitoring data (columns 1 and 4); (2) model estimates for times and locations (counties) with monitoring estimates (columns 2 and 5); and (3) all estimates based on modeling results (columns 3 and 6). The first and third of these summaries use all the monitoring or model simulation results available and thereby have different sample sizes; however, the second ensures an identical sample size in terms of counties and days with data, facilitating comparison of monitor- and model-derived exposure estimates. Model-derived estimates provide greater spatial and temporal coverage than monitor-based estimates. County level monitor-based estimates cover 21.5% (5.63×10^5 km^2) and 26.0% (6.81×10^5 km^2) of land area included in CMAQ exposure estimates (study domain $\sim 2.65 \times 10^6$ km^2) for $PM_{2.5}$ and O_3, respectively. The 2000 population included in the 1861 counties with exposure estimates based on the CMAQ model is 173,675,971. County level exposure estimates based on monitors included 66.5% (population = 115,494,521) of this total population for $PM_{2.5}$ and 67.1% (population = 116,536,577) for O_3. Overall, approximately 23.4% of the population (40,640,177 persons) in the study area resides in a county without either a $PM_{2.5}$ or O_3 monitor. For counties with both monitor- and model-derived exposure estimates, 100% of days had data for the modeling approach: on average, 44.4% and 69.8% of days had data for $PM_{2.5}$ and O_3, respectively, using monitor data (for O_3, 97.9% of days between April–September had data).

Monitor-derived county level annual average $PM_{2.5}$ and seasonal average (April–September) O_3 concentrations are shown in Figs. 4(a) and 5(a), respectively. Corresponding model-derived concentrations are shown in Figs. 4(b) and 5(b). These figures demonstrate differences in spatial coverage between monitor- and model-based

Table 1
Comparison of population characteristics of counties with and without monitors for $PM_{2.5}$ or O_3.

	Median value of census variable (95% Confidence interval) Counties with monitor(s)(n= 412, population = 121.2×10^6)	Counties without monitor(s) (n= 1449, population = 52.5×10^6)
$PM_{2.5}$		
Population Characteristics (% of county population)		
Self-identified as black[a]	16.3 (14.7, 17.8)	12.9 (12.0, 13.8)
Young children ($\leq$5 years)[a]	7.87 (7.77, 7.96)	7.60 (7.54, 7.65)
Elderly ($\geq$65 years)[a]	13.2 (12.8, 13.5)	14.6 (14.4, 14.7)
Urban[a]	71.2 (68.8, 73.6)	32.9 (31.6, 34.2)
High school diploma[a]	32.0 (31.3, 32.6)	37.3 (36.9, 37.6)
Baccalaureate degree [a]	13.6 (13.1, 14.2)	8.86 (8.65, 9.06)
Unemployed[a]	3.60 (3.50, 3.70)	3.36 (3.29, 3.43)
Poverty[a]	13.0 (12.5, 13.5)	14.9 (14.5, 15.3)
Use public transport[a]	2.45 (1.82, 3.07)	0.470 (0.424, 0.516)
Median income ($)[a]	39,786 (38,820, 40,752)	34,152 (33,699, 34,605)
Pollution exposure estimates		
Annual average $PM_{2.5}$ concentration (monitor, μg/m^3)	13.1 (12.9, 13.4)	—
Annual average $PM_{2.5}$ concentration (CMAQ, μg/m^3)[a]	11.6 (10.4, 11.9)	10.2 (10.1, 10.3)

O_3	Median Value of Census Variable (95% Confidence Interval) Counties with monitor(s) (n = 454, population = 116.5×10^6)	Counties without monitor(s) (n = 1407, population = 57.1×10^6)
Population characteristics (% of county population)		
self-identified as black	14.0 (12.6, 15.4)	13.5 (12.6, 14.5)
Young children ($\leq$5 years)[a]	7.91 (7.82, 8.00)	7.58 (7.52, 7.63)
Elderly ($\geq$65 years)[a]	13.0 (12.7, 13.4)	14.7 (14.5, 14.9)
Urban[a]	65.2 (62.6, 67.8)	33.6 (32.3, 35.0)
High school diploma[a]	33.1 (32.5, 33.7)	37.0 (36.7, 37.4)
Baccalaureate degree [a]	13.4 (12.9, 13.9)	8.79 (8.59, 8.99)
Unemployed	3.35 (3.26, 3.45)	3.43 (3.36, 3.50)
Poverty[a]	11.7 (11.2, 12.2)	15.4 (15.0, 15.7)
Use public transport[a]	1.96 (1.50, 2.42)	0.563 (0.443, 0.683)
Median income ($)[a]	41,692 (40,709, 42,674)	33,345 (32,933, 33,757)
Pollution exposure estimates[b]		
Warm season average O_3 concentration (monitor, ppb)	51.9 (51.3, 52.5)	—
Warm season average O_3 concentration (CMAQ, ppb)[a]	53.1 (52.4, 54.1)	52.3 (52.0, 52.5)

[a] Indicates significant differences between groups (counties with and without monitors) at p-value $<$ 0.05.
[b] O_3 pollution exposure estimates are calculated using O_3 season data, i.e., April–September, 2002.

Table 2
Comparison of model result and monitor data and county-level exposure estimates for $PM_{2.5}$ and O_3.

	$PM_{2.5}$			O_3		
	Monitor estimates[a]	Model estimates for locations/times with monitoring data[b]	All model estimates[c]	Monitor estimates[a]	Model estimates for locations/times with monitoring data[b]	All model estimates[c]
Number of observations [d]	73,000	73,000	679,265	116,114	116,114	679,265
Population covered [e]	115,494,521	115,494,521	173,675,971	116,536,577	116,536,577	173,675,971
Area covered, county-level estimates (km^2)	6.31×10^5	$6.31\text{x}10^5$	2.62×10^6	6.86×1000^5	6.86×10^5	2.62×10^6
Overall annual concentration [μg/m^3 for $PM_{2.5}$, ppb for O_3] Mean+standard deviation (minimum to maximum) [f]	13.1+2.02 (7.21 to 12.9)	11.6+2.8 (4.4 to 25.6)	10.3+2.2 (3.2 to 25.5)	47.2+6.5 (29.7 to 60.3) April–Sep. only: 51.9+6.1 (29.3 to 67.8)	49.3+5.2 (26.4 to 59.2) April–Sep. only: 53.1+4.7 (36.9 to 62.6)	44.5+2.8 (25.9 to 50.7) April–Sep. only: 52.4+4.4 (35.8 to 62.5)
Daily concentration [μg/m^3 for $PM_{2.5}$, ppb for O_3] Mean+standard deviation (minimum to maximum) [g]	13.1+7.4 (0.1 to 94.5)	11.6+7.1 (0.04 to 79.4)	10.3+6.2 (0.003 to 91.3)	47.2+18.6 (2.0 to 139.0) April–Sep. only: 51.9+17.5 (3.0 to 139.0)	49.3+14.6 (0.140 to 114.3) April–Sep. only: 53.1+12.7 (0.140 to 114.4)	44.5+13.7 (0.130 to 119.6) April–Sep. only: 52.4+11.5 (0.130 to 119.6)

[a] Includes only monitors meeting inclusion criteria as described in Section 2 (i.e., had sufficient data to be used to generate exposure estimates).
[b] Includes model-derived exposure estimates for counties and times (days) that also have monitor-derived concentration estimates.
[c] Includes all model-derived exposure estimates, irrespective of whether monitor-derived exposure estimates exist at that location and time.
[d] Refers to the total number of county-level exposure estimates derived from monitors and/or the total number of county-level exposure estimates derived from simulation results.
[e] Represents the population residing in the counties for which exposure estimates were generated, using 2000 Census data.
[f] Represents the mean and standard deviation of annual average exposure estimates across all counties, and the highest and lowest annual average exposure estimate for any county.
[g] Represents the mean and standard deviation of daily average exposure estimates across all counties, and the highest and lowest daily average exposure estimate for any county.

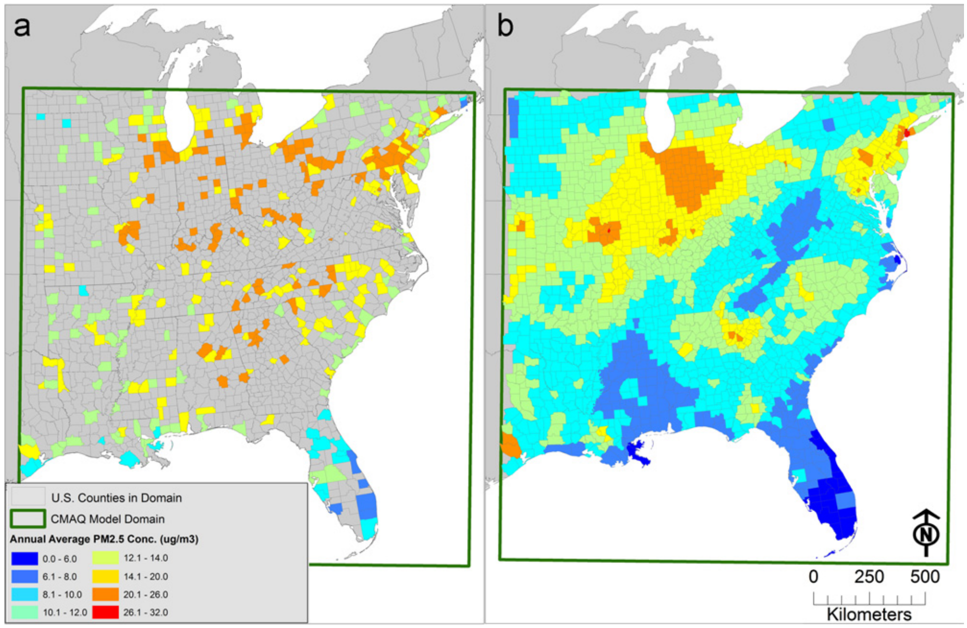

Fig. 4. County-level annual average exposure estimates for 24-hour $PM_{2.5}$ (a) Monitor-derived and (b) model-derived

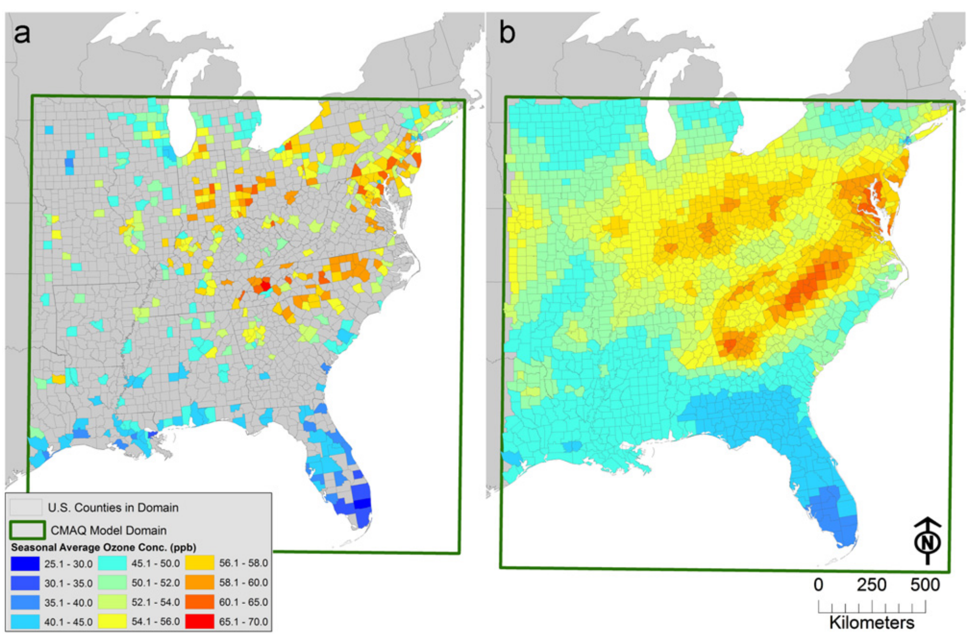

Fig. 5. County-level seasonal average (April–Sept.) exposure estimates for 8-hour O_3: (a) Monitor-derived and (b) model-derived

approaches. For counties with exposure estimates from both approaches, the annual average maximum 8-hour O_3 level was 47.2 ppb for the monitoring approach and 49.3 ppb for the model approach. During April–September, county level O_3 exposure estimates averaged 51.9 ppb using monitors and were somewhat higher (53.1 ppb) using modeling results. County level exposure estimates for $PM_{2.5}$ were 13.1 $\mu g/m^3$ using monitor data, and 11.6 $\mu g/m^3$ using simulation results.

3.3.2. *Location-specific estimates*

Exposure estimates based on specific points were calculated using model and monitor data, similar to how such estimates could be generated for an epidemiological study with information on individuals' locations. Location-specific exposure estimates based on the nearest monitor (within 50 km) results in exposure estimates for 59.8% (1.57×10^6 km^2) and 63.0% (1.65×10^6 km^2) of land area included in model-based exposure estimates for $PM_{2.5}$ and O_3, respectively (Figs. S4 and S5, Supplementary Material). There are significant differences in spatial and temporal coverage between model- and monitor-based approaches. For all locations within the model domain, 100% of days had concentration estimates using model results: on average, 42.6% and 69.9% of days had data for $PM_{2.5}$ and O_3, respectively, using monitor data (for O_3, on average 97.4% of days in April–September had concentration estimates). Typically, epidemiological studies estimating exposures within a certain radius of a given ambient monitor will use a uniform buffer size around monitor locations throughout the study area (Hanigan et al., 2006; Lipsett et al., 2011; O'Donnell et al., 2011; Spencer-Hwang et al., 2011). Using a uniform buffer size has potential to introduce exposure misclassification, particularly for large study areas due to differences in spatial heterogeneity by location. Appropriate buffer size may vary depending on the pollutant of concern, region within the U.S., time of year, population density, long-term ambient pollution concentrations, and other factors (Bell et al., 2011).

4. Discussion

Although reviews of CMAQ performance have identified model strengths and limitations (Baker and Scheff, 2007; Boylan and Russell, 2006; Eder and Yu, 2006; Mueller, 2009; Phillips and Finkelstein, 2006a; Swall and Davis, 2006; Tesche et al., 2006), to the authors' knowledge, no previous work considers how CMAQ performance issues could affect exposure estimates that might be used in epidemiological studies of air pollution and health. Our analysis demonstrates that CMAQ performance is different for $PM_{2.5}$ and O_3, and also depends on the measure used to gauge performance (e.g., bias, error, correlation), season, the time interval for which each metric is calculated, and location within the U.S. In the case of $PM_{2.5}$, long-term averages derived from model results are similar to those derived from monitor data, but short-term averages (e.g., daily, monthly) overestimate observed $PM_{2.5}$ levels during winter and underestimate levels in summer. Annual O_3 levels derived from model results overestimate observed concentrations, but limited monitoring data were available outside of April–September, which hinders assessment of performance during other months. During months when O_3 levels are higher, the model may tend to underestimate O_3 concentrations, evidenced by slightly negative normalized mean bias values for June–September (Fig. 1). In addition, our evaluation of normalized mean bias and correlation by location (e.g., Figs. 2 and 3) indicates that annual average normalized mean bias and correlation vary somewhat by region, with distinct spatial patterns in variation for $PM_{2.5}$ and O_3.

Traditional exposure assessment using ambient monitors excludes populations distant from monitors (although definitions of "distant" vary) (Chen et al., 2007; Jerrett et al., 2004; Sarnat et al., 2006). Based on our exposure estimates and demographic data, approximately 58 and 57 million people in the study area live in counties without $PM_{2.5}$ and O_3 monitors, respectively. Nearly 41 million people live in counties without either type of monitor. We found that populations in counties with and without monitors substantially differed by racial composition, median income, percent of population that are young children or elderly, and levels of poverty, employment, and education. Several studies have indicated that some populations may respond to air pollution differently (Bell and Dominci, 2008; Evans and Kantrowitz, 2002; O'Neill et al., 2003a). Health effect estimates from specific locations, representing certain populations, may not be applicable to the general population (Sarnat et al., 2009). Thus, differences between populations covered and not covered by the monitoring network observed in this study may hinder the ability of epidemiological studies to fully characterize health effects for the general population or to study how demographic factors affect susceptibility to air pollution using observations from ambient monitors. This highlights the need for alternative approaches to exposure assessment.

Addressing bias and errors in simulation results (as compared to monitored observations) can be aided by model calibration, bias correction, and other methods. Studies indicate that bias correction is a useful tool for improving model forecasts of both O_3 and $PM_{2.5}$ concentrations (Delle Monache et al., 2006; Djalalova et al., 2010; McKeen et al., 2005), even across large study areas (e.g., North America, eastern U.S.). Results of a comparison of two bias correction approaches (hybrid filter, Kalman filter) applied to CMAQ simulation results indicate that these techniques reduced systematic errors in model forecasts, although residual error from unsystematic and random errors remained (Kang et al., 2010; Kang et al., 2008). The study also noted that just as model performance varies across space, the efficacy of bias correction techniques exhibited spatial variability, which must be considered with large study areas. Another study evaluated model performance and compared five different bias correction approaches using CMAQ simulation results for New York State and $PM_{2.5}$ and O_3 data from U.S.EPA monitors. Overall, adjusted simulation results were in closer agreement with observed ambient concentrations, but improvements gained through a given bias correction approach tended to differ depending on pollutant, the metric used to measure overall error or reduction in error, and the range and magnitude of ambient concentrations, with some adjustment approaches best at reducing bias at higher observed concentrations (Hogrefe et al., 2006).

While model calibration and bias correction techniques may be useful in improving model forecasts of observed concentrations, these methods also have limitations. Typically, bias correction techniques can only be applied to locations with monitoring data. Further research is needed to develop methods for extending these techniques to areas without monitoring data. Thus, bias correction and model calibration techniques are limited in their ability to address issues such as the lack of data in rural areas, where there is no monitoring data for calibration, or on days without monitoring data (e.g., colder months for O_3 and days throughout the year for $PM_{2.5}$ monitors with one-in-three-day sampling schedules). One key advantage of using regional air quality modeling results to estimate exposure is the ability to estimate exposures and thereby, health effects, in locations and times without monitoring data. Until exposure estimates can be improved, one viable approach to address systematic bias in air quality modeling results is to statistically incorporate the uncertainty into epidemiological analysis.

Furthermore, other efforts are underway to incorporate regional air quality modeling into exposure estimates, including development of approaches combining modeled and measured data (Fuentes and Raftery, 2005; MacMillan et al., 2010). For example, "fused" data uses spatial-temporal Bayesian hierarchical modeling that integrates information from monitoring observations with output from regional air quality models (e.g., CMAQ), to estimate ground-level air pollution concentrations, and has been applied to $PM_{2.5}$ and O_3 (Fuentes, 2009). Such statistical methods are aimed at using multiple types of information to inform exposure estimates, and also allow researchers to estimate exposure in areas far from monitors. However, a limitation of these methods is the introduction of additional uncertainty into resulting exposure estimates. Different spatial resolutions of monitoring data compared to modeling output may introduce bias into pollution or exposure estimates produced by the fused approach, prior distributions used for different parameters in the statistical model may differ by location and air pollutant (Fuentes, 2009; Gotway and Young, 2002), and model performance and accuracy of exposure estimates in locations with little or no monitoring data is difficult to evaluate.

In addition to fused data, a number of other approaches have been developed to estimate individual- and population-level exposures, including various interpolation methods (e.g., kriging), land use regression models, air dispersion and human exposure models, aerosol measurements obtained from satellites, and source- and traffic-proximity analysis (Jerrett et al., 2005a; MacMillan et al., 2010; Nerriere et al., 2005; Paciorek and Liu, 2009; Stein et al., 2007;

Wong et al., 2004; Zou et al., 2009). Interpolation methods (e.g., nearest neighbor, inverse distance weighting, kriging) have been used to estimate air pollution exposures in previous studies (Cohen et al., 2009; Finkelstein et al., 2003; Kunzli et al., 2003), but there is not yet consensus on which methods are most appropriate for estimating ambient concentrations to assess health effects. The quality and certainty of estimated exposures are related to the degree of monitor coverage and spatial heterogeneity of the pollutant within the study area, while the potential for exposure misclassification persists because estimates are based on ambient monitoring data and not personal exposure information (Son et al., 2010; Wong et al., 2004).

Land use regression models utilize information on land use, population density, traffic volume, distance to pollutant source, and ambient pollutant concentrations, and may be able to capture smaller-scale heterogeneity in intra-urban pollutant concentrations (Jerrett et al., 2005a). While this method is transferable to different locations, data availability (e.g., road, traffic, land cover, air pollution monitoring) and quality are potentially significant limitations (Zou et al., 2009). Air dispersion models use information on meteorological conditions, temperature, topography, road type, vehicle speed, emissions, and dispersion processes to estimate pollutant concentration profiles. These models can be applied to different areas or regions of the study area and over different spatial scales (Lipfert et al., 2006), and can provide ambient concentration estimates in locations without dense monitoring networks. However, assumptions must be made regarding the chemical and physical transformation of pollutants and dispersion patterns, model validation is hindered by the limited spatio-temporal resolution of available monitoring data, and model simulations may require significant resources in terms of input data and expertise (Zou et al., 2009).

Exposure and inhalation models using information on ambient pollutant concentrations, human activity patterns (e.g., time spent in microenvironments), physiology (e.g., age, sex), and environmental conditions have also been developed to estimate exposure and health impacts (Fryer et al., 2006). Human inhalation models can model linkages between adverse health outcomes and air pollution and estimate exposures for individuals (Burke et al., 2001; Ozkaynak et al., 2008), but can only be utilized in areas with time–activity data that estimate amount of time spent in different microenvironments (McCurdy et al., 2000).

Remote sensing is yet another method with potential to improve spatial and temporal resolution of measurements of ambient pollutant concentrations. Aerosol optical depth as measured by satellites is correlated with ground level $PM_{2.5}$ concentrations in several studies (Koelemeijer et al., 2006; Liu et al., 2007a, 2007b, 2007c, 2009, 2005; Pelletier et al., 2007). Thus far, aerosol optical depth measurements from satellites have not been used extensively as estimates of exposure to $PM_{2.5}$ in locations without where there is little $PM_{2.5}$ monitoring data for validation (Paciorek and Liu, 2009), although this is an area of active research and improvements.

Lastly, proximity models operate on the assumption that exposure at locations proximate to an emissions source are higher. Utilizing geographic information systems, proximity models may be useful in reducing likelihood of exposure misclassification (Nuckols et al., 2004; Zhan et al., 2006). However, proximity models do not consider pollutant dispersion or human time–activity patterns and may be less appropriate for secondary pollutants and non-traffic related pollutants (Ivy et al., 2008). Research has also suggested that the basic assumption that closer proximity to a source means greater exposure may not always be valid (Cordier et al., 2004).

How CMAQ performance issues affect exposure and health effect estimates depends on the type of epidemiological study. Use of model estimates may introduce differential uncertainty in exposure estimates by season, which is critically important for studies evaluating daily impacts throughout the year. This issue is of particular concern for pollutants such as $PM_{2.5}$, for which health effect estimates vary by season (Bell et al., 2008). High correlations between observed and modeled concentrations indicate that modeled and monitored concentrations tend to increase and decrease in tandem. For time-series and case-crossover studies assessing short-term exposure (e.g., days) and comparing risks across time within a given community, the relative difference between observations and modeled values matched in time and space may be more important than absolute pollution levels (Bell et al., 2004b). In other words, over- (or under-) estimation of pollutant levels by CMAQ may be less problematic for such studies as long as estimated and observed values co-vary in similar patterns.

Cohort designs are suitable for measuring short- and long-term health effects (Kunzli et al., 2001), and compare exposure levels and health response between different populations or communities. For these studies, accurate assessment of differences in pollution/exposure levels between groups being compared is critical. Exposure assessments using air quality modeling results would be hindered by regional variation in model performance, if groups being compared represent different locations and communities. Variation in model performance across time could also detrimentally impact exposure estimates for cohort studies as some parts of the cohort may have better exposure estimates than others.

Our results indicate key strengths of using 3-D air quality models to estimate air pollution exposure in health studies including improvements in spatial and temporal coverage. In this analysis, use of CMAQ simulation results improves sample size, but also changes the nature of scientific questions that can be addressed. For example, daily data is required to perform distributed lag epidemiological models of how health responds to cumulative exposure over previous days. Improved spatial coverage allows study of health effects in rural areas, which may differ with respect to the air pollution mixture, pollution level, or population characteristics. These benefits should be weighed against limitations, such as model performance, the appropriateness of which will depend on epidemiological study design, and the expertise and information required to run CMAQ or similar models.

The CMAQ model is updated and improved as the science advances or if specific issues are identified, and inputs (e.g., emissions inventory data) and precursor models are also revised periodically. The ambient air quality monitoring network also changes over time, as monitors are added or removed, or new monitoring techniques are implemented. The frequent changes in the CMAQ modeling system and input data may affect CMAQ performance in issues critical to use of model results in epidemiological studies. Considering results from this analysis, it may be advisable to conduct a case-specific evaluation of whether a regional air quality simulation is appropriate to use for a given exposure assessment or health study. Air quality modeling is an emerging method for air pollution exposure assessment with some clear advantages over traditional approaches; evaluation of strengths and weaknesses ultimately depends on intended application of model results, acceptable level of uncertainty, population of interest and other factors.

Funding and ethics

This work was supported by U.S. Environmental Protection Agency STAR grant award R833863 (M. Fuentes, Y. Zhang, M. Burr, and M. Bell), U.S. Environmental Protection Agency - Harvard

Clean Air Center grant award RD83479801 (M. Bell and M. Bravo) and National Institute of Environmental Health Sciences grant award R01ES019560 (M. Bell and M. Bravo). The authors declare no conflicts of interest. This work did not involve research on human or animal subjects.

Supplementary material available

Detailed evaluation and demographic data analysis results and additional maps of estimated exposure levels. This material is available free of charge.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2012.04.008.

References

Appel, K.W., et al., 2007. Evaluations of the community multiscale air quality (CMAQ) model version 4.5: Sensitivites impacting model performance, Part I — Ozone. Atmos. Environ. 41, 9603–9615.

Bailey, E.M., et al., 2007. A comparison of the performance of four air quality models for the Southern Oxidants study episode in July 1999. J. Geophys. Res., 112.

Baker, K., Scheff, P., 2007. Photochemical model performance for $PM_{2.5}$, sulfate, nitrate, ammonium, and precursor species SO_2, HNO_3, and NH_3 at background monitor locations in the central and eastern United States. Atmos. Environ. 41, 6185–6195.

Bell, M.L., 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study of ozone in the northern Georgia Region of the United States. Environ. Int. 32, 586–593.

Bell, M.L., Dominci, F., 2008. Effect modification by community characteristics on the short-term effects of ozone exposure and mortality in 98 US communities. Am. J. Epidemiol. 167, 986–997.

Bell, M.L., et al., 2007a. Spatial and temporal variation in $PM_{2.5}$ chemical composition in the United States for health effects studies. Environ. Health Perspect. 115, 989–995.

Bell, M.L., et al., 2007b. Ambient air pollution and low birth weight in Connecticut and Massachusetts. Environ. Health Perspect. 115, 1118–1125.

Bell, M.L., et al., 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. J. Exposure Sci. Environ. Epidemiol. 21, 372–384.

Bell, M.L., et al., 2008. Seasonal and regional short-term effects of fine particles on hospital admissions in 202 U.S. counties, 1999–2005. Am. J. Epidemiol. 168, 1301–1310.

Bell, M.L., et al., 2004a. Ozone and short-term mortality in 95 U.S. urban communities, 1987-2000. J. Am. Med. Assoc. 292, 2372–2378.

Bell, M.L., et al., 2004b. Time series studies of particulate matter. Annu. Rev. Publ. Health 25, 247–280.

Boylan, J.W., Russell, A.G., 2006. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmos. Environ. 40, 4946–4959.

Brauer, M., et al., 2008. A cohort study of traffic-related air pollution impacts on birth outcomes. Environ. Health Perspect. 115, 680–686.

Burke, J.M., et al., 2001. A population exposure model for particulate matter: case study results for $PM_{2.5}$ in Philadelphia, PA. J. Exposure Sci. Environ. Epidemiol. 11, 470–489.

Burr, M.B., Zhang, Y., 2011. Source apportionment of $PM_{2.5}$ over the eastern U.S., Part I. Source sensitivity simulations using CMAQ with the Brute Force Method. Atmos. Pollut. Res. 2, 300–317.

Byun, D., Schere, K., 2006. Review of the governing equations, computational algorithms, and other components of the Models-3 community multiscale air quality (CMAQ) modeling system. Appl. Mech. Rev. 59, 51–77.

Chen, L.P., et al., 2007. Spatiotemporal relationship between particle air pollution and respiratory emergency hospital admissions in Brisbane, Australia. Sci. Total Environ. 373, 57–67.

Cohen, M.A., et al., 2009. Approach to estimating participant pollutant exposures in the multi-ethnic study of atherosclerosis and air pollution (MESA Air). Environ. Sci. Technol. 43, 4687–4693.

Cordier, S., et al., 2004. Risk of congential anomalies in the vicinity of municipal solid waste incinerators. Occupational Environ. Med. 61, 8–15.

Delle Monache, L., et al., 2006. Ozone ensemble forecasts: A Kalman-filter predictor bias correction. J. Geophys. Res., 111.

Djalalova, I., et al., 2010. Ensemble and bias-correction techniques for air qulaity model forecasts of surface O_3 and $PM_{2.5}$ during the TEXAQS-11 experiment of 2006. Atmospherc 44, 455–467.

Dominici, F., et al., 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. J. Am. Med. Assoc. 295, 1127–1134.

Eder, B., et al., 2006. An operational evaluation of the Eta-CMAQ air quality forecast model. Atmos. Environ. 40, 4894–4905.

Eder, B., Yu, S., 2006. A performance evaluation of the 2004 release of Models-3 CMAQ. Atmos. Environ. 40, 4811–4824.

Evans, G.W., Kantrowitz, E., 2002. Socioeconomic status and health: the potential role of environmental risk exposure. Annu. Rev. Publ. Health 23, 303–331.

Finkelstein, M.M., et al., 2003. Relation between income, air pollution, and mortality: a cohort study. Can. Med. J. Assoc. 169, 397–402.

Fryer, M., et al., 2006. Human exposure modeling for chemical risk assessment: a review of current approaches and research and policy implications. Environ. Sci. Policy 9, 261–274.

Fuentes, M., 2009. Statistical issues in health ipact assessment at hte state and local levels. Air Qual. Atmos. Health 2, 47–55.

Fuentes, M., Raftery, A.E., 2005. Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. Biometrics 61, 36–45.

Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. J. Am. Stat. Assoc. 97, 632–648.

Hanigan, I., et al., 2006. A comparison of methods for calculating population exposure estimates of daily weather for health research. Int. J. Health Geogra., 5.

Hogrefe, C., et al., Exploring approaches to integrate observations and CMAQ simulations for improved air quality forecasts. Models-3 Users' Workshop, Chapel Hill, NC, 2006.

Holloman, C.H., et al., 2004. A Bayesian hierarchical approach for relating $PM_{2.5}$ exposure to cardiovascular mortality in North Carolina. Environ. Health Perspect. 112, 1282–1288.

Ivy, D., et al., 2008. Development of ambient air quality population-weighted metrics for use in time-series health studies. J. Air Waste Manage. Assoc. 58, 711–720.

Janes, H., et al., 2007. Trends in air pollution and mortality: an approach ot the assessment of unmeasured confounding. Epidemiology 18, 416–423.

Jerrett, M., et al., 2005a. A review and evaluation of intraurban air pollution exposure models. J. Exposure Sci. Environ. Epidemiol. 15, 185–204.

Jerrett, M., et al., 2004. Do socioeconomic characteristics modify the short term associations between air pollution and mortality? Evidence from a zonal time series in Hamilton, Canada. J. Epidemiol. Commun. Health 58, 31–40.

Jerrett, M., et al., 2005b. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16, 727–736.

Kang, D., et al., 2010. Assessment of bias-adjusted $PM_{2.5}$ air quality forecasts over the continental United States during 2007. Geosci. Model Dev. 3, 309–320.

Kang, D., et al., 2008. Bias adjustment techniques for improving ozone air quality forecasts. J. Geophys. Res., 113.

Koelemeijer, R., et al., 2006. Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. Atmos. Environ. 40, 5304–5315.

Kunzli, N., et al., 2003. Ambient air pollution and atherosclerosis in Los Angeles. Environ. Health Perspect. 113, 201–206.

Kunzli, N., et al., 2001. Assessment of deaths attributable to air pollution: Should we use risk estimates based on time series or on cohort studies? Am. J. Epidemiol. 153, 1050–1055.

Lipfert, F.W., et al., 2006. Traffic density as a surrogate measure of environmental exposures in studies of air pollution healht effects: Long-term mortality in a cohort of U.S. veterans. Atmos. Environ. 40, 154–169.

Lipsett, M.J., et al., 2011. Long-term exposure to air pollution and cardiorespiratory disease in the California teachers cohort study. Am. J. Respir. Crit. Care Med. 184, 828–835.

Liu, Y., et al., 2007a. Using aerosol optical thickness to predict ground-level $PM_{2.5}$ concentrations in the St. Louis area: a comparison between MISR and MODIS. Remote Sens. Environ. 107, 33–44.

Liu, Y., et al., 2007b. Component concentrations and size distributions using satellite retrieved fractional aerosol optical depth: Part I - method development. J. Air Waste Manage. Assoc. 57, 1351–1359.

Liu, Y., et al., 2007c. Estimating $PM_{2.5}$ component concentrations and size distributions using satellite retrieved fractional aerosol optical depth: Part II — A case study. J. Air Waste Manage. Assoc. 57, 1360–1369.

Liu, Y., et al., 2009. Estimating regional, spatial, and temporal variability of $PM_{2.5}$ concentrations using satellite data, meteorology and land use information. Environ. Health Perspect. 117, 886–892.

Liu, Y., et al., 2005. Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing. Environ. Sci. Technol. 39, 3269–3278.

MacMillan, N.J., et al., 2010. Combining numerical model output and particulate data using Bayesian space-time modeling. Environmetrics 21, 48–65.

McCurdy, T., et al., 2000. The national exposure research laboratory's consolidated human activity database. J. Exposure Sci. Environ. Epidemiol. 106, 566–578.

McKeen, S., et al., 2005. Assessment of an ensemble of seven real-time ozone forecasts over eastern North America during the summer of 2004. J. Geophys. Res., 110.

Miller, K.A., et al., 2007. Long-term exposure to air pollution and incidence of cardiovascular events in women. New Engl. J. Med. 365, 447–458.

Morris, R., et al., 2009. Technical Support Document for VISTAS Emissions and Air Quality Modeling to Support Regional Haze State Implementation Plans. Novato, CA.

Mueller, S.F., 2009. Model representation of local air quality characteristics. J. Appl. Meteorol. Climatol. 48, 945–961.

Nerriere, E., et al., 2005. Lung cancer risk assessment in relation with personal exposure to airborne particles in four French metropolitan areas. J. Occup. Environ. Med. 47, 1211–1217.
Nuckols, J.R., et al., 2004. Using geographic information systems for exposure assessment in environmental epidemiology studies. Environ. Health Perspect. 112, 1007–1015.
O'Donnell, M.J., et al., 2011. Fine particulate air pollution ($PM_{2.5}$) and the risk of acute ischemic stroke. Epidemiology 22, 422–431.
O'Neill, M.S., et al., 2003a. Health, wealth, and air pollution: advancing theory and methods. Environ. Health Perspect. 111, 1861–1870.
O'Neill, M.S., et al., 2003b. Health, wealth, and air pollution: advancing theory and methods. Environ. Health Perspect. 111, 1861–1970.
Olerud, D., Sims, A., MM5 2002 modeling in support of VISTAS (visibility improvement state and tribal association of the Southeast). North Carolina division of air quality, Research Triangle Park, NC, 2004.
Ozkaynak, H., et al., 2008. Modeling population exposures to outdoor sources of hazardous air pollutants. J. Exposure Sci. Environ. Epidemiol. 18, 45–58.
Paciorek, C.J., Liu, Y., 2009. Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. Environ. Health Perspect. 117, 904–909.
Pelletier, B., et al., 2007. Retreiving of particulate matter from optical measurements: a semi-parametric approach. J. Geophys. Res., 112.
Peng, R.D., et al., 2008. Coarse particulate matter air pollution and hospital admissions for cardiovascular and respiratory diseases among Medicare patients. J. Am. Med. Assoc. 299, 2172–2179.
Phillips, S., Finkelstein, P., 2006a. Comparison evaluation of two leading photochemical air quality models for particulate matter. Atmos. Environ. 40, 4999–5009.
Phillips, S.B., Finkelstein, P.L., 2006b. Comparison of spatial patterns of pollutant distribution with CMAQ predictions. Atmos. Environ., 40.
Pope, C.A.I., et al., 2009. Fine-particulate air pollution and life expectancy in the United States. New Engl. J. Med., 260.
Pope III, C.A., et al., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. J. Am. Med. Assoc. 287, 1132–1141.
Queen, A., Zhang, Y., 2008. Examining the sensitivity of MM5-CMAQ predictions to explicit microphysics schemes and horizontal grid resolutions, Part III – The impact of horizontal grid resolution. Atmos. Environ. 41, 3869–3881.
Ritz, B., et al., 2002. Ambient air pollution and risk of birth defectse in Southern California. Am. J. Epidemiol. 155, 17–25.
Sarnat, S.E., et al., 2006. Spatial considerations in a study of ambient air pollution and cardiorespiratory emergency department visits. Epidemiology 17, S242.
Sarnat, S.E., et al., 2009. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. J. Exposure Sci. Environ. Epidemiol. 20, 135–146.
Son, J.Y., et al., 2010. Individual exposure to air pollutoin and lung function in Korea: spatial analysis using multiple exposure approaches. Environ. Res. 110, 739–749.
Spencer-Hwang, R., et al., 2011. Ambient air pollutants and risk of fatal coronary heart disease among kidney transplant recipients. Am. J. Kidney Dis.. In press.
Stein, A.F., et al., 2007. Hybrid modeling approach to resolve pollutant concentrations in an urban area. Atmos. Environ. 41, 9410–9426.
Swall, J.L., Davis, J.M., 2006. A Bayesian statistical approach for the evaluation of CMAQ. Atmos. Environ.. 40, 4883–4893.
Tesche, T.W., et al., 2006. CMAQ/CAMx annual 2002 performance evaluation over the eastern U.S. Atmos. Environ. 40, 4906–4919.
U.S.EPA, Air Quality System (AQS). Technology Transfer Network (TTN), Vol. 2010.
Wong, D.W., et al., 2004. Comparison of spatial interpolation methods for the estimation of air quality data. J. Exposure Sci. Environ. Epidemiol. 14, 404–415.
Zhan, F.B., et al., 2006. GIS-EpiLink: a spatial search tool for linking environmental and health data. J. Med. Syst. 30, 405–412.
Zhang, Y., et al., 2006a. A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study episode — Part I: Evaluation protocols, databases, and meteorological predictions. Atmos. Environ. 40, 4825–4838.
Zhang, Y., et al., 2006b. A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study Episode, Part III: Diagnostic and Mechanistic Evaluations. Atmos.c Environ. 40, 4856–4873.
Zhang, Y., et al., 2006c. A comprehensive performance evaluation of MM5-CMAQ for the Summer 1999 Southern Oxidants Study Episode, Part II. Gas and Aerosol Predictions. Atmos. Environ. 40, 4839–4855.
Zou, B., et al., 2009. Air pollution exposure assessment methods utilized in epidemiological studies. J. Environ. Monit. 11, 475–490.